



TAMPERE ECONOMIC WORKING PAPERS
NET SERIES

URBAN GOVERNANCE, COMPETITION, AND WELFARE

Hannu Laurila

Working Paper 29

December 2003

<http://tampub.uta.fi/econet/wp29-2003.pdf>

DEPARTMENT OF ECONOMICS

FIN-33014 UNIVERSITY OF TAMPERE, FINLAND

ISSN 1458-1191
ISBN 951-44-5875-3

URBAN GOVERNANCE, COMPETITION, AND WELFARE

5.5.2004

Hannu Laurila
Department of Economics
FI-33014 University of Tampere

Abstract

The market mechanism of spatial resource allocation is examined in a system of cities, where social welfare depends on city size. The competitive dynamics of the system is a product of the interplay between people's individual exit type choices (migration) and their collective voice type choices (urban governance). It is shown that the use of efficiency enhancing measures of urban governance depends on the pressure of exit. A necessary condition for dynamic efficiency is that the market equilibrium of migration is non-stable, which sounds somewhat paradoxical. Dynamic efficiency is more likely to emerge between initially small cities, in which agglomeration economies dominate, than between initially large cities, in which agglomeration diseconomies dominate. The incentives for proper urban governance are somewhat ambiguous in the most relevant case, where cities are of asymmetric size. It is therefore important to strengthen the incentives by means of national urban policy.

Key words: exit and voice, city size, migration, reactive and proactive policy measures, Tiebout model

JEL classification: 931

An earlier version of the paper is published in:
Urban Studies, Vol. 41, No. 3, March 2004, p. 683-696

1 Introduction

In the past few years internal migration has been lively in Finland as a reflection of globalisation and interEuropean economic integration. The still ongoing structural change from rural economy towards urban economy has been shifting the locational advantages of the Finnish production sector from rural areas to urban areas. At the present stage of development, the focus of the migration pattern has turned from rural-urban migration to inter-city migration. The fact that market forces now determine the location patterns of firms and people more directly than before, and that the elements of welfare nowadays lie in urban surroundings urges the need for economic research concerning the concepts of spatial market adjustment, national urban policy and local urban governance.

People and firms are the key market agents of spatial reformation of the economy. This aspect is well covered by the literature on regional economics and urban economics. The market mechanism is based on welfare and profit maximising choices of location, and on agglomeration economies and diseconomies that connect these choices to community size. However, as collectives of their residents, the communities can also be regarded as market agents in the reformation process. The competitive role of the communities is emphasised in the more recent literature on local public economics.

An early version of a comprehensive treatment that combines, at least implicitly, the above-mentioned elements of migration, optimal community size and competitive governance is the Tiebout model (Tiebout, 1956). The model is a well-known representation of competitive public good provision, or to put it more fundamentally,

of efficient allocation of resources in a spatial context. The model is constructed on two cornerstones, inter-community migration of people seeking better local public services, and competition of the communities for the migrants through attractive tax-service packages. The eventual outcome is a Pareto optimal equilibrium, where nobody can gain by moving to another community.

In the original Tiebout model, the role of competition is sketched very briefly to rest on the implicit analogy between communities and private firms as competitive market agents. The focus is on competition of residents between the communities. The Tiebout model is essentially a demand side model that says little about the supply side, and nothing about the production technology of the localities or the political process inside them (Rubinfeld, 1987, p. 174). This paper attempts to elaborate Tiebout's intuition further by exploring the rationale and preconditions of inter-community competition by turning the focus on the competitive pressures that emerge within the communities. The welfare of the residents in a community is taken to be determined by the size of the community on the one hand, and on the actions of the community itself on the other hand. A simple club theoretic framework consisting of a system of urban areas, henceforth called cities, is used. The model is further extended from that of Tiebout by including the outcomes of private goods and factor markets and the quality of life in people's welfare considerations, and by taking the conditions of the local private markets into the sphere of public influence.

The evolution of the system of cities is studied from the viewpoint of people's choices. Clustering of the firms is an essential source of agglomeration economies, but the firm sector is for simplicity ignored by assuming that the profit maximising

choices of the firms respond competitively to the evolution of the local markets, and that they are effectively reflected by the parameters of the budget constraint faced by the people. People make their choices between and within the cities. Between the cities people make their choices by exit, that is, by voting with their feet between residential sites. Within the cities people make their choices by voice, which is to say that they take democratic and other collective actions so as to affect the local preconditions of their welfare. The collective mode of people's decision-making is the origin of what is here called urban governance. It is also assumed that the preferences of the city residents are effectively revealed on the average by means of local democracy, and that they are carried out undistortedly by the local bureaucracy.

Exit-type choices are close but not equal to normal market choices. In private goods and factor markets the market agents exert exit between preferred and non-preferred items, and they choose optimal amounts of the preferred ones. Residential choices are different, because people are not able to carry out the second stage of decision-making – they can not optimise on the amount that is the size of the city, which indirectly determines their utility in the city. Furthermore, exit-type choices are made according to average, not marginal concepts, as they are made in the market case. Therefore, exit does not suffice to secure efficient market outcomes in a general sense (Atkinson & Stiglitz, 1982). However, by raising their voice the residents of a city are collectively able to optimise on the size of the city. In collective decision-making, the problem of marginal vs. average choice criteria is also irrelevant (Cornes & Sandler, 1998).

The necessary condition of allocative efficiency is that all market agents operate in a technically efficient manner. As is well established in the standard theory of the firm,

competition is the main motor of technical efficiency. Competitive dynamics in the system of cities precludes external pressure to evoke technical efficiency within the cities as in the case of competing firms. At first glance the analogy seems to be straightforward - the threat of emigration (exit) should force the cities to efficient conduct of urban governance (voice) just as firms have to respond to declining market shares with better efficiency. The situation, however, is more complicated. In the following chapters the nature of competitive dynamics of the system of cities is examined in close detail. An analogy between the position of private firms and that of the cities is sought.

2 The basic model

2.1 Agglomeration economies and welfare

A conventional postulation in the literature of urban economics is that welfare in a city depends on its size (Alonso, 1971). This is because household, business and social agglomeration economies (and diseconomies) that arise in densely populated urban surroundings determine both the benefits and the costs of urban life (Richardson, 1979, p. 306). This classification points to the fact that locally experienced welfare depends on the efficiency of the activities of the firms and the local public sector in the city. In this paper, the setting is simplified by assuming that the firm sector is competitive and adjusts efficiently to the respective circumstances, and only the private and collective actions of the households are studied.

The optimisation problem of a site-seeking household can be stated as a standard consumer's choice by assuming that the household agglomeration economies are experienced directly on the benefit side (in the utility function), and that the business and social agglomeration economies are faced indirectly on the cost side (in the budget constraint) (compare to Richardson, 1973, pp. 11-20). The direct benefits of urban location are anticipated in everyday life, and the arguments of the utility function include quantities of private goods, public goods and leisure. The quality of life can also be included in the leisure variable. Agglomeration economies enter as technical externalities. The costs of urban life, described by the budget constraint, are determined by local factor prices, particularly wages on the income side, and by commodity prices, transportation costs, local taxes and tariffs etc. on the expenditure side. Here agglomeration economies work indirectly through the price mechanism, and are thus pecuniary in nature. Because of agglomeration economies, both the benefits and the costs are implicitly determined by the size of the city. (Mills & Hamilton, 1984; McCann, 2001.)

In the spirit of constrained utility maximization the welfare anticipated by an average citizen is determined by the net sum of benefits and costs of urban life. A standard assumption is that agglomeration economies dominate the formation of welfare in the earlier stages of city growth, but that agglomeration diseconomies gain strength with growth and eventually become dominant (Richardson, 1973). Figure 1 illustrates the basic assumptions.

(Figure 1 about here)

In Figure 1, city size is measured by the number of people (denoted by n) within the geographical area of the city. This is reasonable if the geography is taken as given in the short term. In the figure two sets of curves are presented. The curve $AW^p = W^p/n$, where W^p is potential total welfare, plots the potential average welfare consisting of the net sum of benefits and costs in the city considered. The term potential refers to full utilisation of the existing welfare creation ‘technology’ of a city. The technology itself is taken as given and innovations that change it are ignored. In this sense, the AW^p curve resembles the curve of the attainable profits of a competitive firm, or the production possibility frontier of an economy. The curve $MW^p = dW^p/dn$ plots the potential marginal net benefits that new migrants bring to total welfare in the city. The MW^p curve strikes through the AW^p curve from above at n^o , where the latter reaches its maximum point. At that point, agglomeration diseconomies start to dominate agglomeration economies.

The other set of curves in Figure 1, namely AW and MW , plots the actually observed average welfare in the city. Due to market distortions and excess transaction costs in the local private sector and inefficiencies in the local public sector, the actual observations may lie far below the welfare potential of the city (Brueckner, 1982). This reflects inefficient utilisation of the existing welfare creation technology. The observation of welfare is the only relevant information. In Figure 1, at city size n^l , the average citizen observes welfare AW_l , while the theoretical potential would be AW^p_l . Denoting the gap between potential and actual welfare by g , the total welfare worth of g times n^l remains unexperienced in the city. This is the dead weight loss of remaining below the potential of the city. Furthermore, $n^p > n^o$ would be the optimal population if the full potential were reached.

Since both the benefit side and the cost side are determined by the size of the locality, the number of the population should be a true choice variable in the optimisation of the households, but it is evident that the households cannot choose optimal population such as n^o or n^p in Figure 1. Therefore, the maximisation problem for an average household stated above actually is the problem of a social planner. For this reason (among other things) the city must now also be studied from the point of view of collective choice.

2.2 *The role of urban governance*

Urban governance concerns a broad range of public activities that affect the preconditions of welfare in the city (Bailey, 1999). The essential feature here is that urban governance is based on collective decision-making - the citizens make their collective choices by voice, that is, via local democracy. A benchmark case for efficient urban governance is the following: the distribution of welfare in the city is symmetrical, the local public expenditures are financed by lump sum taxes, and the local bureaucracy operates without bias. Then, the welfare of the median voter represents average welfare, and a simple majority-voting rule produces a Pareto-optimal outcome. It is obvious that these features are quite specific in nature. (Cullis & Jones, 1998, pp. 78-87.)

Urban governance can be divided into reactive and proactive functions. In the reactive sense urban governance is operated on a short-term basis with the prevailing circumstances taken as given. The focus is on the observed changes in the average welfare - systematic migration into the city may increase or decrease the welfare of

the existing residents. From this perspective, the main function of urban governance is the optimization of population. In Figure 1 population n^o is chosen along the AW schedule as a true short-term solution to the consumer's problem depicted above.

There are some limits to the horizontal optimisation of population. Aiming at n^o from left to right is out of compass, because that depends on the exogenous exit decisions of potential newcomers. Aiming at n^o from right to left may also be difficult to implement, at least if the civil and property rights of the current residents are respected. The function can be operated in practice only by allowing entrance to the city up to n^o (or other stipulated population level) and closing entrance at that size. The operative measures include planning, zoning, housing policy and dimensioning of public services. The reactive optimisation of population is broadly discussed in the literature on urban economics and it is also in common use in practice. The optimisation criteria, however, are often other (say, real estate values) than that used in this paper. (Brueckner, 1982.)

The proactive function of urban governance concerns full utilisation of the welfare potential of the city. In terms of Figure 1 this refers to bridging the gap between potential and actually observed welfare by shifting the observation AW_I vertically closer to the potential AW^p_I , and refining the shape of the AW curve to follow that of the AW^p curve. Proactive urban governance takes care of the technical efficiency of the city as an economic unit. Just as in the case of profit maximising private firms in competitive markets the survival of a city depends on its technical efficiency. The proactive functions of urban governance include competition policy, structural policy and growth policy.

The concept of competition policy refers here to improvements in the efficiency of the local public and private sectors. Recalling the above benchmark for efficient urban governance, efficiency in the public sector must be secured through decision-making, financing and execution of service production. It is well stated in the literature of collective decision-making that in practice Pareto-efficient outcomes are seldom reached by majority voting. The decision-making methods must be developed to ensure that urban governance is properly steered by the preferences of local residents. The promises of new technology have received attention in this respect. In public finance, distortive proportional and progressive taxes are much more familiar in practice than undistortive lump sum taxes. Refinements of tax systems, applications of marginal cost pricing and charging policies are examples of financial reforms. The executive problems of bureaucracy are well known. Standard means of improving the efficiency of service provision include privatisation, monitoring, deregulation, competitive bidding, purchaser-provider arrangements etc. (Bailey, 1999.)

Proactive urban governance also concerns the efficiency of the local private sector because it has a vital role in the functioning of the local commodity and factor markets (Prud'homme & Lee, 1999). Urban governance can notably affect both the market distortions faced by the competitive firm sector and the overall transaction costs in the city. As a key constructor of local technical and social infrastructure, urban governance can also contribute to the formation of local social capital, which has a catalyst role in motoring up innovation, and a pre-emptive role in establishing a fruitful economic environment (Kajanoja & Simpura, 2000). On the technical side, the instruments include zoning and city planning and systems of public transport, logistics

and networks, which have considerable effects on the time and money costs of market transactions in the city. As to social infrastructure, the instruments include immaterial networks and recreational, cultural and other such facilities.

Structural policy means motoring up and co-ordinating reallocation of local resources from sunset industries to sunrise ones. The main instrument in this respect is strategic programming. Traditional measures include focusing of local budget funds and external funding in order to evoke economic incentives for warranted development. More modern modes of local structural policy aim to amalgamate the mutual interests of the community and local private business life into joint efforts of strategic development (Bailey, 1999).

Growth policy is about setting long-term size (or growth rate) targets. As it is the density of population that determines agglomeration economies and welfare, both the geographical area and population of the city must be controlled. Regarding geographical optimisation, recall that the AW schedules in Figure 1 are drawn for a given area. For areas of different sizes the curves may look different. If economies/diseconomies of scale exist, the AW curve can be made to reach a higher peak value by widening/narrowing the city area. The area must be chosen so that the long-run average welfare is maximised. Instruments of geographical optimisation include municipal consolidations, federations, and regional co-operation in service production and development. Long-term optimisation of population is then about choosing welfare-maximising population along the proper AW^p schedule. The same operational notions apply as in the reactive mode of urban governance, and the

instruments are mostly the same, too. As is demonstrated later, proper conduct of this function is vital for the efficiency of the whole system of cities.

To conclude, voice-driven urban governance is operated under the pressure of exit-type migration decisions. Reactive optimisation of population is a response to inward migration. It is easy to defend the welfare of the original citizens by stopping planning for new housing and services just before the requested city size is reached. The decision is politically neutral, at least if the preferences of the median voter equal those of the average household. In the median voter model the outcome is also welfare maximizing in the short term. Reactive urban governance resembles firms' short-term optimisation on production according to market parameters. The effort of applying the proactive functions needs pressure from the threat of outward migration. This is because implementation of these functions affects the present residents of the city. Equity considerations arise because more efficient tax and transfer policies treat people unequally, structural reallocation of resources causes unemployment in some industries, revisions in city planning and public transport systems capitalize in real estate values, and so on. The likelihood of political conflicts is high, and the costs of collective decision-making increase (Cullis & Jones, 198, pp. 82-83). There is again an analogy to private firms: the powerful example of dynamic inefficiency of a monopoly suggests that such of internal renewals are accepted only under considerable market pressure (Bailey, 1999, pp. 98-100). Let us now make the pressure of inter-city migration exit explicit in the analysis.

3 Inter-city migration and competitive dynamics

3.1 Large cities

Assume first that the economy consists of a system of initially large and homogeneous cities. Let the number of cities be m , and let the total population be N . People are perfectly mobile and fully informed about the welfare observations everywhere, and migration between cities emerges if welfare differentials are found. This is a standard assumption, and the introduction of mobility costs and uncertainty in welfare comparisons would not change the basic intuition. Figure 2 below illustrates this standard setting. In the figure, the MW curves as well as the AW^p and MW^p curves are ignored for simplicity.

(Figure 2 about here)

In Figure 2 two of the identical cities, city A and city B, are presented. The curves AW_i , $i = A, B$, plot average welfare in the cities, drawn from left to right for city A and from right to left for city B. Since all cities are identical the length of the horizontal axis is the share of the two particular cities A and B of the whole population N , namely $n = n_A + n_B = 2N/m$, where n_A and n_B denote the number of residents in A and B respectively. The households make their choices according to observations along the AW_i curves.

Assume that the total population is initially divided so that the number of people in A and B are $O_A n^1$ and $O_B n^1$ respectively. At this stage, there exists an observed welfare

gap, denoted by g , in favour of city A. People are induced to move from city B to city A. Migration narrows the gap because of two effects. First, migration into A enforces agglomeration diseconomies in A, which lowers benefits and raises costs causing average welfare in A to fall. Second, migration from city B decreases agglomeration diseconomies in B, which increases welfare for those who remain in B. Continued migration leads to a stable equilibrium at point e , where welfare in both cities is equalised and no motives for further migration exist. In the equilibrium, the size of city A is $n_A = O_A n^e$ and the size of city B is $n_B = O_B n^e$. It is evident from the figure and from the above definition of n that $n_A = n_B = N/m$. The equilibrium is also efficient, because social welfare, consisting of the area $O_A A W_A^e A W_B^e O_B$ is at its maximum. For any other value of n there would be welfare losses. Note that in this individualistic way of welfare measurement the problem of inter-city externalities and other reasons for paternalistic welfare judgements are ignored (Boadway & Bruce, 1984).

However, the market solution of Figure 2 is efficient only in the static sense. Since the AW_i curves in the figure present observed welfare, assumed to remain below the respective potentials, welfare improvement through higher technical efficiency is possible. The question now is whether there is such a competitive pressure in the system that would provoke improvements in technical efficiency in the cities as there is in the case of private firms operating in competitive markets. The competitive pressure needed for improvements in efficiency is generated by exit. Since the equilibrium at e is a stable position with zero welfare differentials, there exists no such pressure at that point. People are not on the move and urban governance is not forced into inconvenient improvements in technical efficiency. The case of a stable market equilibrium does not seem to imply high efficiency in the dynamic sense.

In order to gain deeper insight into the dynamic properties of a stable market solution it is useful to examine an alternative equilibrium path recalling the option of static urban governance. Reconsider the starting point n^l in Figure 2, where migration from B to A makes those who remain in B better off but those who originally live in A worse off. Because the initial residents of A anticipate that in-migration will reduce their welfare, they may be induced to raise their voice at that moment. The simplest thing they can do to avoid the welfare loss is to adopt reactive urban governance and close entrance to city A at n^l . Now, migration is unable to level up the welfare gap g . For city B, the gap can be bridged only by the means of proactive urban governance. The residents in city B may raise their voice in favour of developing the use of welfare creating technology. This is illustrated by the upward shift of the AW_B curve to AW'_B . It must be noted, however, that city B is not forced to do this – there is no threat of exit because people cannot migrate to city A. Another notable implication is that there is no great urgency for improvements in city A even if the average welfare in B should increase above that of A at n^l . This is because migration to B would benefit both A and B.

If city B manages to improve its technical efficiency, the welfare of the whole system increases somewhat. This is because B moves closer to its potential, and the original residents of B become better off. However, the situation is still sub-optimal if no improvements in technical efficiency in city A have occurred. On the other hand, if city B is not able to improve its efficiency, the welfare of the whole system remains lower than in the market solution at point e . The general conclusion is that the equilibrium at point e' is basically stable just like that at point e , and no exit-driven

competitive pressure exists to draw the AW_i curves closer to their potential counterparts.

3.2 Small cities

The above analysis confirms the standard result that with free migration there is a case for a stable market equilibrium. However, because of low incentives for improvements in technical efficiency this equilibrium may well remain dynamically sub-optimal. The result holds in the case of large cities, which operate on the falling regimes of their average (and marginal) welfare schedules. It is worthwhile also to analyse the case in which the cities in the system are initially small and agglomeration economies dominate. This is also the situation that is more relevant in practice, at least in a national perspective. Figure 3 depicts this case.

(Figure 3 about here)

In Figure 3 the setting is changed from that of Figure 2 by assuming that the system now consists of a considerably larger number of cities $m' > m$, while the total population N is taken to be unchanged. Therefore, in any representative city, the number of residents must now be smaller, or, to put it in other words, the horizontal axis must be shorter with respect to the AW (and MW) schedules.

Reconsider the situation of welfare equalisation presented by point e in Figure 3.

With small cities, point e is evidently not a stable market equilibrium. Suppose that, for purely stochastic reasons, some households move from city A to city B, and that the allocation of people shifts from n^e to n^l . A welfare gap g^0 is opened in favour of

B. In this case, migration does not level up the gap - in fact it draws the gap even wider and makes the solution to deviate from e at an accelerating rate. The only stable market equilibrium in this case would be a corner solution, in which either city A or city B, depending on the direction of the initial stochastic shock, is deserted.

In the non-stable migration pattern reactive urban governance is of no use. Closing entrance does not help to stop out-migration from city A, and city B does not want to cut the beneficial in-migration. The incentive structure of urban governance is now changed from that in Figure 2. In Figure 3, migration from city A to city B makes the old and new residents of city B better off because of rising agglomeration economies, but the residents remaining in city A become worse off because of decreasing agglomeration economies. The residents of the out-migration city now suffer, and thus face the pressure of exit. Proactive urban governance is to be considered.

Under the pressure of accelerating exit by their fellow citizen, the more reluctantly moving residents of city A raise their voice in favour of competition policy and structural policy. Figure 3 illustrates how improvements in local efficiency shift average welfare in A up to AW'_A so as to bridge the welfare gap g^0 . At e^1 , average welfare in both cities is again equalised. Assume that city A now looks even more attractive compared to city B at the marginal, and migration turns backwards resulting in e^2 . A welfare gap g^1 is now opened in favour of city A. People in city B have to respond to this by expressing their desire for efficiency improvements, which results in an upward shift of AW_B to AW'_B . If migrants are now attracted to city B, people in city A will again be forced to achieve further improvements. The process continues in

consecutive order. As a result of proactive urban governance in both cities, the actually observed welfare levels converge to their potential counterparts.

In principle, the competition for migrants between A and B will continue as long as welfare gains can be created by improvements in technical efficiency. In the long-term all cities will go through the evolution illustrated in Figure 3. After reaching their potential welfare curves, the cities can no longer compete for new residents by improving technical efficiency. At this stage, the non-stable nature of the solution appears confusing, and a reasonable question is whether an efficient market outcome is at all possible in the long run.

Proactive growth policy now comes into the picture. All cities optimise on their geographical area and choose their population so as to maximise the average welfare in that area (Ng, 1973). Those cities that manage to attract people allow migrants to enter until average welfare in the city reaches its maximum. After that, entrance is closed. The remaining cities will continue the competition. Some of them attract people and some lose them. The number of cities is endogenously determined in the long run. The mechanism continues until there is an optimal number of cities all having an optimal population. The cities may turn out to be heterogeneous with homogeneous population in them. Ignoring the integer problem, all people in every city enjoy maximal welfare and nobody wants to migrate. The solution is socially Pareto efficient since nobody's welfare could be improved without making somebody else worse off. (Cornes & Sandler, 1986.) The famous rank-size rule (or Zipf's law) of 'natural' size distribution of cities can be regarded to reflect this kind of development (McCann, 2001, p. 79).

The above general equilibrium solution, of course, remains hypothetical and the social costs of the evolution process may be significant in practice. The dismal fact is that some cities are doomed to be deserted and at least some infrastructure must be relocated as far as total population is taken as fixed. The problem of depopulation can be avoided by taking total population endogenous. The rural sector or other countries, ignored in the above analysis, can to some extent serve as a population reserve in this respect. But if there is not such an effective reserve, social costs may arise. If the indivisible costs of demographic and economic decline fall mainly on the residents of the declining cities, the migration rate may be biased to exceed that of the 'natural' rate of depreciation of the infrastructure. The focus regarding the problem of sunk costs is on the speed of the adjustment.

3.3 *Asymmetric cities*

So far the discussion has focused on symmetric situations along falling or rising average welfare schedules in both cities alike. In practice, there are often cities in both positions at the same time. Another rather restrictive assumption made above is the assumption of homogeneity of the welfare schedules between cities. Due to various geographical, historical, institutional, structural and other reasons, the welfare schedules of the cities may well be different from each other even in the longer term. The economic preconditions of the cities are quite unilateral, especially in recently industrialized countries, like Finland. For practical reasons, it is most reasonable to take these asymmetries into account in the analysis.

Figure 4 below illustrates the essentials of the case of asymmetric cities. In the figure, the AW curves for cities A and B are again presented, but city A is now assumed to be initially a more fertile city with higher average welfare around the peak value than in city B. The AW curves are drawn to reflect the difference. The AW curves intersect at two points, namely at e and e' in order to demonstrate all the remaining variations of possible outcomes.

(Figure 4 about here)

At the intersection point e in Figure 4, city A is initially large as compared to city B. City A is on its falling welfare regime due to high agglomeration diseconomies, whereas agglomeration economies dominate in city B. Just as in the case of large cities above, it is obvious that point e is basically a stable market equilibrium. Consider first allocations right from e . There is a welfare gap attracting people from A to B. The adjustment will result in a stable market solution at e . Since migration is beneficial to both A and B, and the adjustment levels up the welfare gap, no notable incentives for raising voice exist in either city. As to allocations left from e , say at n^l , there is a welfare gap that draws people from B to A towards a stable solution at point e . This would be inconvenient for both A and B. If city A now closes entrance at n^l , city B can respond by the effort of proactive urban governance. If the effort is sufficient to level up the welfare gap so that AW'_B is reached, the system ends up at a stable equilibrium at n^l . It must be noted that residential reallocations leftwards from e would improve social welfare even without dynamic improvements in B. The incentive structure reduces to that of the case of large cities. Incentives for dynamic improvements in efficiency are weak around the stable intersection point e .

At the other end in Figure 4, point e' presents a more promising case, where city B is large and city A is small. Agglomeration economies dominate in A and diseconomies dominate in B. The solution at point e' is non-stable, since starting from e' in either direction opens up exploding welfare gaps. If, for stochastic reasons, migration should start leftwards from e' , both cities would lose welfare. In this direction, the large city B can and most probably will stop the loss of its residents' welfare by closing entrance. With entrance to B closed, city A may try to level up the welfare gap by proactive urban governance, but again it is not the pressure of exit that necessitates this. The possibility of closing entrance makes e' stable leftwards.

If migration should turn rightwards from e' in Figure 4, that is, if people should move from B to A, both cities would gain welfare. City A gains because of increased agglomeration economies and B gains because of decreased diseconomies. The case can be interpreted as one in which a small modern city attracts people from a larger city with a less developed industrial structure. Because there is now an ever-expanding welfare gap, the point e' is unstable in nature. In city A there certainly is no apparent need for raising voice, but in city B the incentives for urban governance are somewhat more ambiguous. The incentives rise because of the accelerating out-migration and expanding welfare gap, but the incentives are dampened by the fact that the welfare of the still remaining citizens is boosted in any case.

To get an idea of the dynamics, suppose that migration from B to A continues as such rightwards from e' until the allocation n_A^* is reached. At this moment the need for urban governance becomes apparent in city A: city A adopts the reactive mode and

closes entrance to optimize on its population. This eliminates the pressure of exit, but leaves a wide welfare gap open. Suppose also that even without immediate pressure of exit city B succeeds in implementing proactive urban governance and manages to level up the welfare gap at point e'' in Figure 4. At this stage, the competitive dynamics of the system depends on the shape of the average welfare schedule of city B. In Figure 4, the upwards-shifted average welfare curve AW_B'' has a positive slope at n_A^* . This is to say that city B has managed to implement a major reform regarding the effects of agglomeration economies. In this case the solution e'' is non-stable, and the system moves to a dynamically efficient general equilibrium just as in the case of small cities presented above. This is because, left from point e'' , both cities would gain welfare from in-migration and are therefore forced to improve their technical efficiency to attract the migrants. But, if the AW_B'' schedule would slope downwards at n_A^* (as presented by the broken line version of the average welfare curve), there is no threat of exit, and point e'' remains a stable and dynamically sub-optimal solution.

To sum up the findings, competitive incentives for improving technical efficiency are generally rather weak in the case of asymmetric cities. The incentives may remain too shallow to force the costly and inconvenient measures of proactive urban governance to be applied. Dynamic efficiency is likely to occur only in the case of a genuinely non-stable market solution. Such a situation is possible as a special case, where there is a welfare gap in favour of the initially small city (rightwards from point e' in Figure 4). However, dynamic efficiency rests on two rather restrictive preconditions. First, the large out-migration city should improve its technical efficiency even though it gains welfare from the outflow (moving from e' to e'' in Figure 4). Second, the improvement in technical efficiency must be substantial enough to make the effects of

agglomeration economies more positive at the margin than those in the rival city (the average welfare curve must be upwards sloping at point e'' in Figure 4). This is to say that the optimal size of city B should be far beyond its short-term optimum n_B^* .

The above observation stresses the importance of national level policy: the shallowness of the incentive structure must be enforced. Furthermore, it is not only that the larger cities must be encouraged to use proactive urban governance, but the focus must be placed on heavy structural reforms that motor up the effects of agglomeration economies. In the case of an initially large city the necessary reform sounds major but not impossible considering the disposable resources. The large cities must be motivated to grow even larger beyond their short-term optimality considerations. This benefits not only the cities themselves but also the entire society, because it facilitates competitive dynamics and, consequently, dynamic efficiency in the city system.

4 Conclusions

The analysis complements the message of the Tiebout model of spatial resource allocation by investigating the preconditions of competitive dynamics in a system of cities. In the present model agglomeration economies explain the connection between city size and welfare, and local welfare is maximised by urban governance (voice) under the market pressure created by migration (exit). The dynamic efficiency of the spatial system rests on the interplay of exit and voice. The incentives for dynamic improvements depend on the competitive nature of the system. In this respect, the cases of small and asymmetric cities are insightful.

Urban governance is divided into reactive and proactive functions. The reactive function is short-term optimisation of population, operated by closing entrance into the city. This function resembles the profit maximising production decisions of firms. It can be implemented in a simple manner without notable political conflicts.

Proactive urban governance concerns long-term improvement in the utilisation of the welfare-creating technology of the city. This function is analogous to the reorganisation of a firm's operative practices in utilising the available production technology. Implementation of this function necessitates considerable effort and is due to evoke inconvenient equity and other such considerations. Therefore, it is undertaken only under competitive pressure.

A somewhat paradoxical finding is that dynamic efficiency in a system of cities precludes non-stable migration patterns. In stable market situations the competitive dynamics remain partial and sub-optimality of resource allocation results in dynamic sense, because there is no pressure of exit to compel the cities to use proactive urban governance. Stability depends on the initial stage of utilisation of agglomeration economies. If all cities are initially large, meaning that agglomeration diseconomies dominate in all cities, the market solution is stable in nature. This is also true in the case of asymmetric cities, particularly if the larger cities implement reactive optimisation of population. The market situation is clearly non-stable if all cities are initially small, i.e. that agglomeration economies dominate everywhere, and both the cities that gain and those that lose population are forced to make improvements in their technical efficiency. Non-stable market situations also exist in an important special case, where small modern cities draw people from larger traditional cities, in

which agglomeration diseconomies dominate. In these two cases there is a close enough analogy to private firms operating in competitive markets.

Reactive urban governance is a major cause of stability, because closing entrance into in-migration cities eliminates the pressure of exit in the out-migration cities. If population is optimised with reference to average welfare below the potential capacity of the city, the outcome is socially sub-optimal in the dynamic sense, and in most cases also in the static sense. Optimisation of population is socially beneficial only when connected to long-term growth policy. Optimal population targets should therefore be set only with reference to long-term welfare potentials to ensure the efficiency of the long-term general equilibrium in the city system. It is clear, however, that this function remains utterly theoretical in practice. From the social point of view, optimisation of population must be considered very critically.

The analysis provides a clear-cut principle for national urban policy. On the one hand urban policy must work indirectly to enhance the working of the market mechanism. In this respect the pressure of exit must be strengthened and proper conduct of voice must be encouraged. The pressure of exit can be enforced by social and institutional reforms that enhance the mobility of people and remove distortions from the price mechanism that steers residential choices. To enable better conduct of voice the autonomous position of the cities as true market agents must be strengthened with particular emphasis on the implementation of proactive instruments of urban governance. Sufficient fiscal and administrative autonomy must be assigned to economically arranged urban areas. The Finnish system of highly autonomous local governments is basically a solid ground for such development. The urban areas must

be made to understand that they do not compete with smaller places and rural areas – they compete with each other for the welfare increasing migrants, and the competition is becoming more and more international in nature. Therefore, the welfare potential and growth of the largest cities is of particular importance. Short-term optimisation of population must be prohibited and the cities must be motivated to concentrate on proactive urban governance. Harsh structural take-offs must be encouraged especially in large cities with a traditional and possibly outdated industrial structure, and which seem to be reluctant to adopt newcomers.

On the other hand national urban policy has direct functions. The market-oriented mechanism cannot handle important issues such as inter-regional externalities and equity. Traditionally the negative externalities have been of major concern, but nowadays the positive externalities attached e.g. to education, research, health care, and to the formation on social capital in general have received more and more attention. Furthermore, as was indicated above, proactive urban governance quite evidently raises up considerations of vertical equity. Horizontal equity among cities and regions, which has traditionally been a central element in the Finnish welfare state, can be endangered, too. All these issues must be dealt with in the framework of fiscal federalism, and clearly national level functions have to be compassed directly by means of urban policy.

The above conclusions are of particular interest in Finland. In the present stage of globalisation and economic integration, the Finnish economy must compete against the circumstances on the European single market. The former comparative advantages, largely attached to the rural sector, have shifted to the new absolute

advantages of the urban economy. The important role of agglomeration economies has become evident, which is also reflected in the recent trend towards concentration. Moreover, the undeniable handicap caused by geography and climate stresses the fact that spatial factors carry more and more weight in the determination of absolute advantages. In a northern country, the costs of housing and transportation are high compared to European rivals. Therefore, the traditional element of agglomeration economies, that is the economics of scale and scope in constructing the inter- and intra-urban infrastructure, is still of particular importance. The fact that the development of the Finnish urban sector must be promoted has also been recognised in the recently introduced national urban policy. The policy is constructed to rest on the intrinsic development of economically determined urban areas, and aims at deeper integration of the communities in those areas. So far there has been insufficient political courage to focus the national urban policy on the issue of city size.

In the light of the above analysis, the issue of city size, measured both in terms of population and geographic area, seems to be of immense importance to Finnish urban policy and urban governance. Finnish cities vary in size and in their economic, historic, geographic and other such determinants, but nevertheless they are all small and newborn in international comparison, and are scattered over a large country. Even the most rapidly growing urban areas are geographically wide and their population densities remain low.

From the above perspective, at least four questions for further research can be formulated. First, there is need to evaluate the present practices of urban policy and urban governance. The Finnish urban policy quite ambitiously aims at developing a

wide range of very broadly defined urban areas, and the policy of some central cities seems to be quite negative regarding the adoption of newcomers. It could be argued that a higher degree of concentration would be needed to yield significant agglomeration economies. The assignment of responsibilities between national urban policy and local urban governance must be clarified. Second, the positive feature that the Finnish cities are comparatively autonomous fiscally and administratively does not necessarily concern the functional urban areas. Means of integration of the municipalities in the urban areas should be investigated. Third, it would be worthwhile to explore the true potentials of the larger cities by national and international comparisons e.g. by means of data envelopment analysis. It is a quite reasonable assumption that there must be plenty of room for improvement in their technical efficiency. And fourth, the effects of proactive urban policy should be evaluated. As to competition policy, the emphasis should be on efficiency improvements in the field of the local private markets. The effects and means of structural and growth policy should be studied with particular interest.

References:

ALONSO, W. (1971) The Economics of Urban Size. **Papers and Proceedings of the Regional Science Association**, 26.

ATKINSON, A. and STIGLITZ, J. (1980) **Lectures on Public Economics**. London: McGraw-Hill Book Company Ltd.

BAILEY, S. (1999) **Local Government Economics. Principles and Practice**. Houndmills: MacMillan Press Ltd.

BOADWAY, R. and BRUCE, N. (1984) **Welfare Economics**. Oxford: Basil Blackwell Ltd.

BRUECKNER, J. (1982) A Test for Allocative Efficiency in the local Public Sector, **Journal of Public Economics**, 19.

CORNES, R., and SANDLER, T. (1986) **The Theory of Externalities, Public Goods, and Club Goods**. Cambridge: Cambridge University Press.

CULLIS, J. and JONES, P. (1998) **Public Finance and Public Choice**. New York: Oxford University Press Inc.

KAJANOJA, J. and SIMPURA, J.(eds.) (2000) **Social Capital, Global and Local Perspectives**. VATT-Publications 29. Helsinki: Government Institute for Economic Research.

LAURILA, H. (2000) Urban Governance – kaupunkitalouden tehokas hallinta, **Kunnallistieteellinen aikakauskirja**, 1/2000.

McCANN, P. (2001) **Urban and Regional Economics**. Oxford: Oxford University Press.

MILLS, E., and HAMILTON, B. (1984) **Urban Economics**. New York: Harper Collins College Publishers.

NG, Y. (1973) The Economic Theory of Clubs: Pareto Optimality Conditions, **Economica**, vol. 4, no. 159, August.

RICHARDSON, H. (1973) **The Economics of Urban Size**. Westmead: Saxon House/Lexington Books.

RUBINFELD, D. (1987) **The Economics of the Local Public Sector**. In: Handbook of Public Economics, vol. II, A. Auerbach and M. Feldstein (eds.). Amsterdam: Elsevier Science Publishers B.V. (North Holland).

PRUD'HOMME, R. and LEE, C. (1999) Size, Sprawl, Speed and the Efficiency of Cities, **Urban Studies**, vol.35, no. 11.

TIEBOUT, C. (1956) A Pure Theory of Local Expenditures, **Journal of Political Economy**, 64, October.

Figure 1: Potential and actual welfare in a city

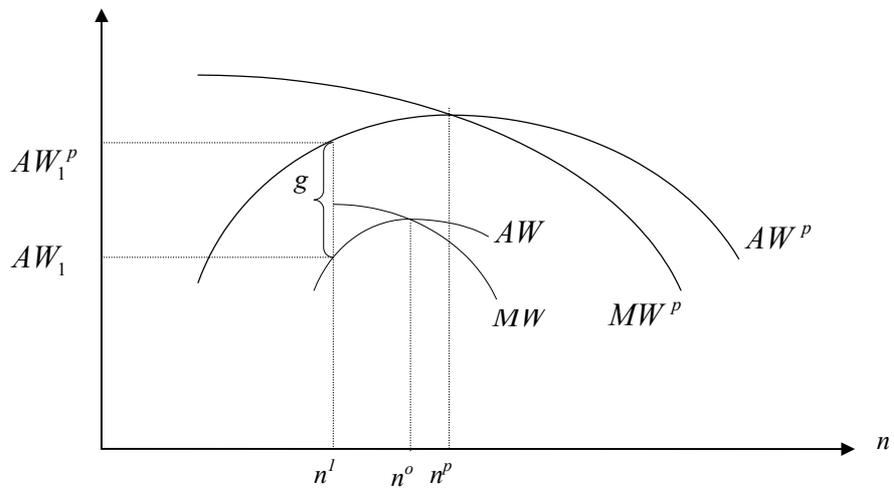


Figure 2: Market equilibrium with large cities

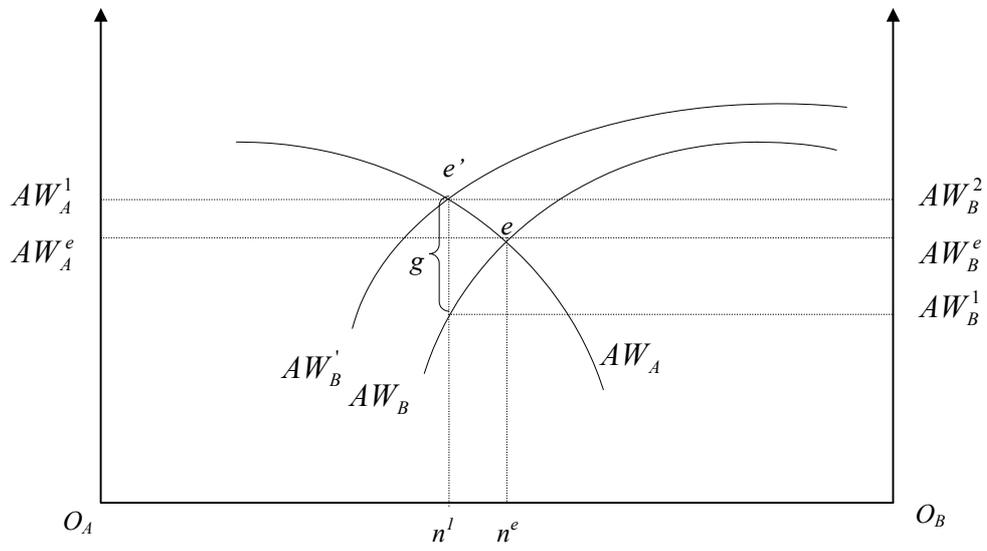


Figure 3: Competitive dynamics in the case of small cities

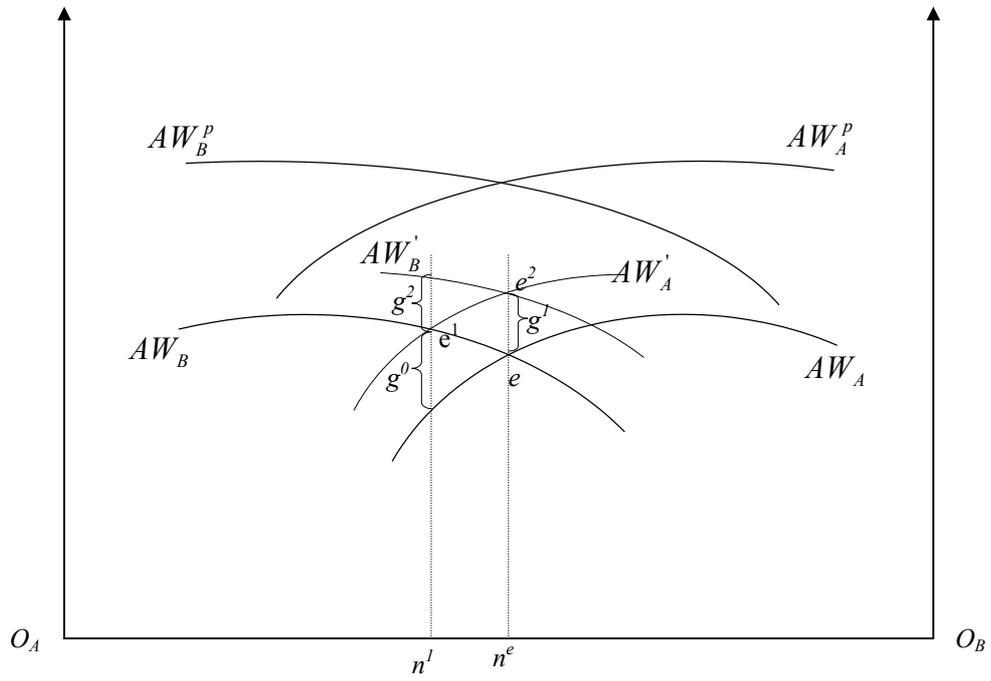


Figure 4: The case of asymmetric welfare potentials

